

# Die glorreichen sieben Datenarten

Ein Dilemma dieser Tage ist der Spagat zwischen der anwachsenden Informationsflut und der gleichzeitig zunehmenden Anforderung, diese Flut auch in geregelte Bahnen zu lenken. Aller Anfang ist schwer, was jedoch hilft, ist eine Klassierung der Daten auf Unternehmensebene.



**«Das Raster der sieben Datenarten ist die Basis für einen gezielten Umgang mit der Datenflut.»**

Daniel Liebhart ist Dozent für Informatik an der Hochschule für Technik in Zürich und Solution Manager der Trivadis AG.

Seit Jahren beobachten Institute wie die University of California, die Data Mobility Group oder auch das Internet Systems Consortium, was wir in der täglichen Arbeit erleiden: die Zunahme der Daten. Sowohl die Speicherung als auch der Transport der Daten hat längst den Exabyte-Bereich erreicht. Unternehmen sind gefordert, aus diesen Daten die wesentlichen Informationen zu extrahieren, und sie sind gefordert, die Vielzahl der Daten durch eine vernünftige Gesamtstrategie so in den Griff zu bekommen, dass sie zur richtigen Zeit in der richtigen Form am richtigen Ort verfügbar sind. Und, dass die gesetzlichen Regelungen erfüllt werden und die Risiken, Daten an Unbefugte zu verlieren, minimiert werden können.

## Was hilft

Um die grosse Menge und zunehmende Anzahl der Daten in einem Unternehmen in den Griff zu bekommen, sind heute viele Bestrebungen im Gange. Das reicht von der Etablierung eines Corporate Data Quality Management über die unternehmensweite Bewertung von Daten bis hin zur Realisierung von speziellen Systemen für Stammdaten oder den Datenaustausch. Die Schwierigkeit liegt immer bei der Zuordnung der verschiedenen Datenarten. Da genügt eine einfache Unterscheidung zwischen strukturierten und unstrukturierten Daten einfach nicht mehr. Auch eine Unterscheidung zwischen zustandsorientierten und ereignisorientier-

ten Daten hilft nicht viel. Eine gute Unterscheidung lässt sich anhand der Dimensionen inhaltliche Komplexität und der Anzahl der Werte aufzeigen, und da kommen dann tatsächlich sieben verschiedene Datenarten zum Vorschein.

## Die sieben verschiedenen Datenarten

Die verschiedenen Datenarten sind Metadaten, Referenzdaten, unternehmensweite Strukturdaten, Transaktionsstrukturdaten, Inventardaten, Transaktionsdaten und Auditdaten.

Metadaten sind Daten, die Daten beschreiben. Das typische Beispiel für Metadaten ist der Systemkatalog einer Datenbank, der die Definitionen aller Tabellen und Spalten enthält.

Referenzdaten sind Daten, die generelle Geschäftsentitäten beschreiben oder kategorisieren, beispielsweise Länder- oder Flughafenkürzel, Währungsbezeichnungen oder andere standardisierte Kodierungen. Referenzdaten können in vielen Branchen von externen Quellen bezogen werden.

Unternehmensweite Strukturdaten repräsentieren den Aufbau und die Organisation des Unternehmens selbst und des Leistungsangebots. Also der Produkte und Services und der Zuständigkeiten.

Transaktionsstrukturdaten beschreiben den Aufbau und die Grundstruktur der Transaktionen des Unternehmens. Also typischerweise die an einer Transaktion beteiligten

Geschäftsentitäten wie beispielsweise der Kunde und das Produkt.

Inventardaten sind diejenigen Daten, die Unternehmenswerte und deren Quantität beschreiben. Beispielsweise die Lagerbestände, die Kontostände oder die Anzahl Liegenschaften einer Firma.

Transaktionsdaten werden durch die operativen Systeme erzeugt. Sie beschreiben die Geschäftsaktivität eines Unternehmens, Beispiele sind Verkäufe, Schadensfälle oder Zahlungen.

Die Prüfung der einzelnen Schritte einer Transaktion kann nur dann erfolgen, wenn diese Schritte auch protokolliert werden. Dies geschieht durch Auditdaten, die von jeder Transaktion erzeugt werden. Beispielsweise Datenbank- und Prozesslogs oder auch Weblogs sind Auditdaten.

## Und jetzt?

Das Raster der sieben Datenarten ist der Ausgangspunkt für einen gezielten Umgang mit der Datenflut. Eine Zuordnung sämtlicher in einem Unternehmen vorhandener Datenarten erlaubt sowohl die Realisierung entsprechender Informationssysteme für die Verwaltung bestimmter Datenarten als auch die Etablierung vernünftiger Governance-Strukturen. So kann beispielsweise ein Unternehmen festlegen, dass Referenzdaten, unternehmensweite Strukturdaten und Transaktionsstrukturdaten in einem Master-Data-Management-System zu verwalten sind. Oder die Art und Weise festlegen, wie geschäftskritische Systeme Auditdaten zu liefern haben und wie diese zu speichern sind. Aber auch eine Vielzahl anderer Anwendungen ist denkbar. <